DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Dissociating Confidence and Accuracy: Functional Magnetic Resonance Imaging Shows Origins of the Subjective Memory Experience

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation	Chua, Elizabeth A., Erin Rand-Giovannetti, Daniel L. Schacter, Marilyn S. Albert, Reisa A. Sperling. 2004. Dissociating confidence and accuracy: Functional magnetic resonance imaging shows origins of the subjective memory experience. Journal of Cognitive Neuroscience 16(7): 1131-1142.
Published Version	doi:10.1162/0898929041920568
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:3627270
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Dissociating Confidence and Accuracy: Functional Magnetic Resonance Imaging Shows Origins of the Subjective Memory Experience

Elizabeth F. Chua[1,2], Erin Rand-Giovannetti[3], Daniel L. Schacter[2], Marilyn S. Albert[3,4], and Reisa A. Sperling[1,3]

Abstract

■ Successful memory typically implies both objective accuracy and subjective confidence, but there are instances when confidence and accuracy diverge. This dissociation suggests that there may be distinct neural patterns of activation related to confidence and accuracy. We used event-related functional magnetic resonance imaging to study the encoding of novel face–name associations, assessed with a postscan memory test that included objective measures of accuracy and subjective measures of confidence. We showed specific neural activity in the left inferior prefrontal cortex associated with trials when subjects expressed high confidence that they had chosen the correct name for the face and made a correct identification. Moreover, we found that this region was also associated with imparting high confidence when subjects chose the incorrect name. However, medial temporal lobe regions showed activity only for high-confidence correct trials. Many functional magnetic resonance imaging studies have shown that the medial temporal lobe and left prefrontal regions are particularly important for the successful formation of memories by using a combination of subjective and objective measures. Our findings suggest that these regions may be differentially involved in the objective and subjective components of memory and that the origins of confidence–accuracy dissociations may be related to incomplete activation of the neural pattern seen in successful encoding. These findings may also aid understanding of eyewitness misidentifications and memory distortions. ■

INTRODUCTION

Episodic memory can be assessed in terms of objective accuracy and subjective confidence. Many functional magnetic resonance imaging (fMRI) studies have relied on the combination of objective accuracy and subjective confidence to determine the neural correlates of successful memory formation (Sperling et al., 2003; Reber et al., 2002; Strange, Otten, Josephs, Rugg, & Dolan, 2002; Otten & Rugg, 2001; Kirchhoff, Wagner, Maril, & Stern, 2000; Brewer, Zhao, Desmond, Glover, & Gabrieli, 1998; Wagner et al., 1998). For example, Wagner et al. (1998) found that activation of left inferior frontal and medial temporal regions during encoding was greater for items subsequently remembered with high confidence than for subsequently forgotten items, which were accompanied by lower levels of subjective confidence. Although confidence often weakly predicts accuracy across subjects, there are specific instances when confidence and accuracy diverge, and individuals report high subjective confidence despite inaccurate memory content (Sporer, Penrod, Read, & Cutler, 1995). This dissociation suggests that separable patterns of neural activity may be related to objective accuracy versus subjective confidence. One way to examine the neural underpinnings of the confidence–accuracy dissociation is to examine specific neural events during encoding that promote subsequent confidence, even when subsequent memory is inaccurate. In this study, we used event-related fMRI to investigate the neural correlates of confidence and accuracy for recognition memory, with particular interest in the encoding origins of incorrect confident recognition.

The dissociation between confidence and accuracy, most apparent and potentially most concerning in cases of eyewitness misidentifications, has generated much research (Weber & Brewer, 2003; Bradfield, Wells, & Olson, 2002; Brewer, Keast, & Rishworth, 2002; Wells, Olson, & Charman, 2002, 2003; Wells & Olson, 2003; Perfect, Hollins, & Hunt, 2000; Tomes & Katz, 2000; Bornstein & Zickafoose, 1999; Lindsay, Read, & Sharma, 1998; Sporer et al., 1995). Overall, reports from the eyewitness literature show that confidence and accuracy are generally only weakly related, such that higher

[1]Brigham and Women's Hospital, [2]Harvard University, [3]Massachusetts General Hospital, [4]Johns Hopkins University School of Medicine

© 2004 Massachusetts Institute of Technology

confidence is only somewhat predictive of accuracy across subjects (for a meta-analysis, see Sporer et al., 1995), yet eyewitnesses who express high confidence in their identifications are very convincing to jurors (Wells et al., 2002; Wells & Olson, 2003). The presentation of misinformation (Belli, Lindsay, Gales, & McCarthy, 1994), confirming feedback (Bradfield et al., 2002), and retrieval effort (Shaw & Zerr, 2003) all tend to increase confidence, but not accuracy for an event and may explain why the relationship between confidence and accuracy is weak in the eyewitness literature.

However, eyewitness scenarios are not the only situations in which confidence–accuracy dissociations may occur. Laboratory manipulations have also shown dissociations between subjective confidence and objective accuracy. Busey, Tunnicliff, Loftus, and Loftus (2000) showed subjects pictures of ''bright'' and ''dim'' faces at study and test in a crossover design. They demonstrated that manipulating luminance affected confidence and accuracy in different ways, that is, subjects were more confident about brighter faces overall, and when faces were studied as dim and then tested as bright accuracy was lowered and confidence was increased.

Various theories of recognition memory offer different explanations about how subjective confidence decisions are made. Signal detection theories posit that confidence decisions are made the same way old–new decisions are made and that a different criterion for each confidence decision lies somewhere on the decision axis and is based on item strength. Analyses by Stretch and Wixted (1998) have suggested that these criteria change as accuracy changes. Specifically, they showed that as discriminability decreases, the confidence criteria fan out along the decision axis. In contrast, according to attribution theory, the subjective experience of remembering is based on attributing an experience to the past and that differs from a memory trace (Jacoby, Kelley, & Dywan, 1989). These theories are grounded in recognition processes, rather than encoding. However, examining encoding processes may further our understanding of factors that contribute to the subjective component of memory.

Behavioral studies have shown that subjective confidence for identifications may have some origins in encoding. Olsson and Juslin (1999) found that subjects who rated their encoding of a perpetrator as ''holistic'' (deep) rather than ''analytic'' (shallow) showed higher identification accuracy and a stronger correlation between confidence and accuracy. Frost and Weaver (1997) showed that memory distortions from misleading information could be eliminated with a longer encoding condition and more retrieval cues. These behavioral studies have shown that encoding conditions influence confidence, but the neural correlates of the encoding origins of confidence remain unclear.

Recent developments in event-related fMRI (Dale & Buckner, 1997) have been useful for investigating the neural correlates of encoding. The ''subsequent memory'' paradigm, exemplified by Wagner et al.'s (1998) study cited earlier, allows a trial-by-trial comparison for stimuli that were successfully remembered to those that were forgotten. To distinguish between remembered and forgotten stimuli, most studies have used a postscan recognition test that incorporates both objective and subjective measures of memory. Postscan tests typically measure accuracy with old/new recognition judgments. Additionally, these tests usually include a subjective component. Many studies have required subjects to provide ''remember'' and ''know'' ratings (Henson, Rugg, Shallice, Josephs, & Dolan, 1999; Brewer et al., 1998) or to provide ''high-confidence'' and ''low-confidence'' ratings for the stimuli (Kirchhoff et al., 2000; Wagner et al., 1998). Both of these procedures are used to distinguish memories of different strengths, but may not be equivalent. The remember/know procedure is often used to distinguish different states of retrieval based on recollection and familiarity, and the high/low-confidence procedure is often used to distinguish ''real'' memories and eliminate guesses. Using a combination of objective and subjective measures, studies have shown that prefrontal and medial temporal lobe (MTL) structures are related to successful memory formation (Sperling et al., 2003; Reber et al., 2002; Strange et al., 2002; Otten & Rugg, 2001; Kirchhoff et al., 2000; Brewer et al., 1998; Wagner et al., 1998). However, it remains unknown whether these structures are differentially involved in the objective and subjective consequences of episodic encoding.

Prefrontal and MTL regions may have dissociable roles in the objective and subjective components of memory. The MTL includes the hippocampus and parahippocampal gyrus. Based on a wide body of evidence from animal and human studies, it is believed that the role of the hippocampus and related structures is to bind together distributed representations that constitute a memory (for a review, see Squire, 1992). This idea implies that the hippocampus is primarily involved in the objective component of memory, since its primary role appears to be in forming accurate representations. It is less clear which brain regions are involved in the subjective component of memory. Left prefrontal activation has been reported in a variety of encoding tasks (Daselaar, Veltman, Rombouts, Raaijmakers, & Jonker, 2003), and evidence from studies on monitoring and familiarity have suggested that the prefrontal cortex may be involved in the subjective component of memory. One recent fMRI study during encoding of words showed left inferior prefrontal activation related to subsequent recollection and familiarity; familiarity was measured using confidence decisions (Ranganath et al., 2004). Additionally, recent fMRI studies during retrieval have suggested that the prefrontal cortex is involved in memory monitoring (Dobbins, Foley, Schacter, & Wagner, 2002; Henson, Rugg, Shallice, & Dolan, 2000).

In a previous study, using a blocked fMRI design, we showed that encoding of face–name associations reliably activates the hippocampus, prefrontal, and fusiform cortices (Sperling et al., 2001). Using an event-related ''subsequent memory'' paradigm during the encoding of face–name associations, we then demonstrated that activation in the left inferior prefrontal cortex and bilateral anterior MTL (including the hippocampus proper and the entorhinal cortex) was related to successful associative memory (i.e., there was greater activation in these regions for face–name pairs that were subsequently remembered with high confidence compared with all incorrect responses, which included responses made with both high and low confidence) (Sperling et al., 2003).

In the current study, we used the same event-related fMRI face–name paradigm (Sperling et al., 2003) to focus on the neural correlates of subsequent confidence and subsequent accuracy, with particular attention to incorrect confident recognition. Subjects were scanned while viewing 455 unfamiliar faces paired with fictional first names and were instructed to try to remember the name associated with the face for later testing. At the same time, subjects were also instructed to make a purely subjective decision about whether the name ''fit'' the face. Approximately 15 min after the scanning sessions, subjects were given a forced-choice recognition test. Each face was paired with the correct name and an incorrect name (which was also seen during encoding, but was paired with a different face). They indicated which name was correct and also judged their confidence level (high or low) that they had chosen the correct name. To investigate the neural correlates of the subjective and objective components of memory, we analyzed the fMRI data collected during encoding based on the results of the postscan memory test. Objective accuracy was measured solely in terms of correctly identifying the name associated with the face and subjective confidence was based on the subjects' ratings. There were four different response types for the encoded face–name pairs: high confidence correct (HC-correct), low confidence correct (LC-correct), high confidence incorrect (HC-incorrect), and low confidence incorrect (LC-incorrect). To analyze the subjective component of memory, we compared all high-confidence responses with all low-confidence responses (all HC vs. all LC), high confidence to low confidence within correct responses (HC-correct vs. LC-correct), and high confidence to low confidence within incorrect responses (HC-incorrect vs. LC-incorrect). To analyze the objective component of memory, we compared all correct versus all incorrect, HC-correct versus HC-incorrect, and LC-correct versus LC-incorrect. We also analyzed the interaction between confidence and accuracy. We hypothesized that the regions identified in successful associative memory, namely, the left inferior prefrontal cortex and bilateral MTL regions, would demonstrate dissociable roles in subsequent confidence and accuracy. We predicted that the left inferior prefrontal cortex would be involved in high confidence, and expected it to be activated in all three of the high-confidence to low-confidence comparisons, regardless of accuracy. Conversely, we predicted that MTL regions would be mainly involved in successfully binding previously unrelated items of information, and expected these areas to be active only during the comparison of high and low confidence within correct responses.

RESULTS

Behavioral Results

All 16 subjects that participated in this study were included in all analyses. Chi-square analyses on individual subjects showed that each subject performed significantly above chance levels ($p < .001$) on the postscan forced-choice face–name recognition task. On average, there were 148 ± 37 trials in which subjects had high confidence that they had chosen the correct name for the face and were indeed correct (HC-correct), 46 ± 27 trials in which subjects had high confidence and chose the incorrect name (HC-incorrect), 142 ± 39 trials in which subjects indicated low confidence for the correct name (LC-correct), and 103 ± 22 trials in which subjects had low confidence and chose the incorrect name (LC-incorrect) (Table 1). Each subject had at least 10 trials of each response type (HC-correct, LC-correct, HC-incorrect, and LC-incorrect).

A 2 × 2 repeated measures factorial ANOVA showed significant differences in the number of trials based on response type. Subjects subsequently indicated the correct name significantly more than the incorrect name, $F(1,15) = 218.19$, $p < .001$, but there was no significant main effect of confidence on number of trials. There was, however, also a significant behavioral interaction between confidence and accuracy, $F(1,15) = 55.25$, $p < .001$. Paired t tests showed no significant difference in the number of trials for HC-correct and LC-correct, but there were significantly fewer HC-incorrect than LC-incorrect responses, $t(15) = 4.85$, $p < .001$. There were significantly more HC-correct responses than HC-incorrect responses, $t(15) = 17.03$, $p < .001$, and significantly more LC-correct responses than LC-incorrect responses, $t(15) = 5.88$, $p < .001$.

Table 1. Mean (*SD*) Number of Responses Given in Each Response Type per Subject

	High Confidence	*Low Confidence*
Correct	148 (37)	142 (39)
Incorrect	46 (27)*	103 (22)

*$p < .00001$.

Table 2. Mean (*SD*) Reaction Time (sec) during Encoding for Each Response Type per Subject

	High Confidence	*Low Confidence*
Correct	1.09 (0.15)	1.08 (0.15)
Incorrect	1.10 (0.15)	1.09 (0.14)

Table 2 shows the mean reaction time during encoding for each of the subsequent response types. Repeated measures ANOVAs showed no significant main effects, interactions, or simple effects of confidence or accuracy on response time during encoding.

During encoding, subjects made a subjective decision about whether the name ''fit'' the face to aid in the associative encoding process. *t* tests showed that subjects were more likely to indicate that the name ''fit'' the face than that it did not ($p < .00001$). However, there were no significant differences within each response type in the percentage of face–name pairs that were classified as ''fits'' compared with ''not-fits'' (e.g., the number of HC-correct responses that were ''fits'' out of the total number of ''fits'' compared with the number of HC-correct responses that were ''not-fits'' out of the total number of ''not-fits'').

There was no effect of stimulus order within or across runs on the percentage of correctly (all correct or HC-correct) or incorrectly remembered face–name pairs. There may have been a slight subsequent memory advantage for trials with longer poststimulus fixation periods because of a possible longer encoding period. Fixation time after each face–name pair and overall correct responses were not significantly correlated, but longer fixation periods after each face–name pair did show an extremely weak, yet significant, correlation with the number of subjects who made HC-correct responses ($R = .123$, $p < .01$). In the postscan test, subjects were no more or less likely to choose the correct name associated with a face when the name had been presented previously as a distracter name. However, subjects were less likely to indicate high confidence in their decision when the name had been presented previously ($p < .00001$).

Imaging Results

Subjective Confidence

We used three comparisons of interest to examine the regions activated during encoding of face–name associations that were subsequently linked with high confidence for recognition memory: all HC versus all LC, HC-incorrect versus LC-incorrect, and HC-correct versus LC-correct. The contrast all HC > all LC showed a similar pattern of activation to our previous studies of attempted associative encoding (Sperling et al., 2001), engaging bilateral dorsolateral prefrontal cortex (DLPFC; right = 48, 18, 24; left = −39, 15, 27; $p < .001$), left ventrolateral prefrontal cortex (VLPFC; −39, 27, −6), bilateral MTL (right = 33, −15, −12; left = −27, −15, −12; $p < .001$), and bilateral fusiform regions (right = 39, −42, −12; left = −27, −42, −21; $p < .001$). DLPFC activation was defined as occurring in Brodmann's areas 9/46 and VLPFC activation occurred in areas 45, 47, and 12. The MTL activation was centered in the hippocampus and extended to the entorhinal cortex. However, given that within high-confidence responses there were more correct responses than incorrect responses, it seems likely that this pattern of activation was largely driven by activation associated with high-confidence correct responses. There were no regions that showed significant differences in activity for all HC < all LC.

In our analyses within correct responses, the contrast HC-correct > LC-correct showed greater activation in bilateral DLPFC, left VLPFC, right MTL (which included the hippocampus proper and the entorhinal cortex), and right fusiform regions, whereas the HC-incorrect > LC-incorrect comparison revealed activation only in the left DLPFC (Figure 1; Table 3). To minimize our type II error, we further examined HC-incorrect > LC-incorrect, and neither lowering the extent threshold nor lowering

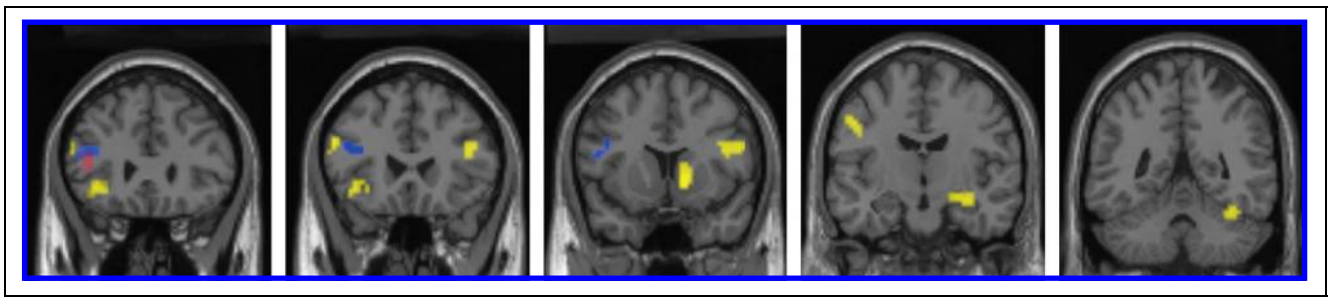

Figure 1. Frontal regions are associated with subsequent high confidence. Group-averaged activation maps for HC-correct > LC-correct (yellow), HC-incorrect > LC-incorrect (red) and area of overlap between these two contrasts (blue) are shown on coronal slices from a single subject template. The left inferior frontal gyrus was activated for both correct and incorrect responses (blue), suggesting that it has a general role in high confidence. HC-incorrect (red) responses showed additional activation in more of the left inferior frontal gyrus and HC-correct (yellow) responses showed activation extending to more superior regions bilaterally, the right hippocampus, the right caudate, and the right fusiform gyrus.

Table 3. Significantly Activated Regions (Presented in MNI Coordinates) for High Confidence Responses With and Without Accuracy

Contrast	*Region*	*x*	*y*	*z*	*Z Score*	*No. of Voxels*
HC-correct vs. LC-correct	Left inferior frontal gyrus	−36	27	−3	3.70	35
	Left middle frontal gyrus	−54	24	24	3.83	23
	Right inferior frontal gyrus	45	18	21	4.15	102
	Right caudate	12	9	0	4.26	85
	Left precentral gyrus	−51	−6	39	4.14	43
	Right hippocampus	33	−12	−15	3.77	50
	Right fusiform gyrus	36	−48	−24	3.61	39
HC-incorrect vs. LC-incorrect	Left inferior frontal gyrus	−42	33	9	3.89	23
''Conjunction'' analyses for HC vs. LC	Left inferior frontal gyrus	−42	24	21	N/A	60

the significance threshold to $p < .01$ showed any significant differences in activation in the MTL. We also used a masking approach (Cabeza, Dolcos, Graham, & Nyberg, 2002) to determine which brain regions demonstrated overlapping activation for the contrasts HC-correct > LC-correct and HC-incorrect > LC-incorrect (Figure 1). The only region that was significantly active in both HC-incorrect > LC-incorrect and HC-correct > LC-correct was in the left DLPFC (Table 3). There were no regions that showed significant differences in activation for HC-correct < LC-correct. One region on the border of the superior and middle frontal gyrus showed significant differences in HC-incorrect < LC-incorrect, but this region did not show significant differences in HC-correct < LC-correct.

To further examine the possibility that the MTL and left prefrontal regions were differentially involved in confidence and accuracy, we extracted the magnetic resonance (MR) signal using a region-of-interest (ROI) approach. The left DLPFC ROI was functionally defined from the overlapping region for HC-correct > LC-correct and HC-incorrect > LC-incorrect contrasts, and the MTL region was functionally defined from the HC-correct > LC-correct contrast. The MR signal time course in the left DLPFC cortex (−42, 24, 21) showed that at the approximate peak of the hemodynamic response there was a significant main effect of confidence, that is, all HC > all LC, $F(1,15) = 10.21, p < .006$. There were also simple main effects of high confidence, that is, HC-correct was greater than LC-correct, $t(1,15) = 2.48, p < .025$, and HC-incorrect was greater than LC-incorrect, $t(1,15) = 2.55, p < .022$ (Figure 2). Within the right MTL (33, −15, −30, shown in Figure 2), at the approximate peak of the hemodynamic response, there was greater MR signal response for HC-correct compared with LC-correct, $t(1,15) = 2.774$, $p < .014$, but there was no significant difference for HC-incorrect > LC-incorrect. It is of note that there was also a significantly greater MR response for HC-correct > LC-incorrect, $t(1,15) = 2.364$, $p < .032$, and that the difference between HC-correct > HC-incorrect approached significance, $t(1,15) = 1.678$, $p < .14$.

Objective Accuracy

We also used three comparisons of interest to examine activation subsequently linked with accuracy for choosing the name associated with the face: all correct versus all incorrect, HC-correct versus HC-incorrect, and LC-correct versus LC-incorrect. At our threshold of $p < .001$ (uncorrected), there were no significant areas of activation for the main effect of accuracy (all correct vs. all incorrect). This null result was not surprising, given that LC-correct responses are more likely to be contaminated by correct guesses, and thus not included in standard subsequent memory analyses (Sperling et al., 2003). However, it is noteworthy that exploratory analyses at a lower threshold of $p < .01$ (uncorrected) showed activation in bilateral MTL regions (right = 27, −15, −24; left = −27, −24, −24; $p < .01$) and in the left frontal operculum (−48, 30, 0), for all correct > all incorrect. There were no regions that showed significant differences in activation for all correct < all incorrect ($p < .001$). For the HC-correct > HC-incorrect contrast, there was significant activation in the thalamus and in the superior frontal gyrus ($p < .001$). Because of our a priori hypotheses about MTL involvement in accuracy for identifying the name associated with the face, we performed exploratory analyses for the HC-correct > HC-incorrect contrast and demonstrated activation in the MTL at a threshold of ($p < .01$, uncorrected). No regions showed significant differences in activation for HC-correct < HC-incorrect. There were no significant differences between LC-correct and LC-incorrect responses in either direction, lending further support to

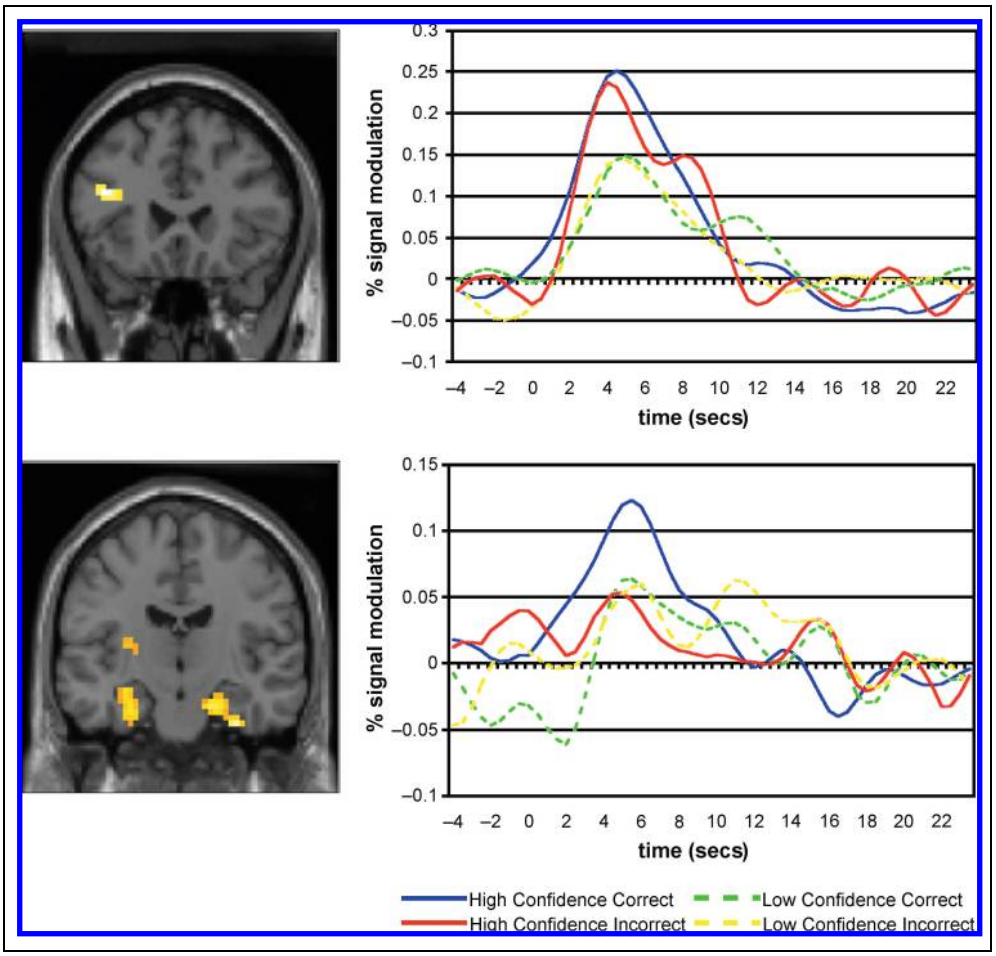

Figure 2. The left inferior prefrontal cortex and right MTL are differentially involved in high confidence. MR signal time courses extracted from the left inferior prefrontal cortex (top) showed that both HC-correct and HC-incorrect responses were significantly greater than low-confidence responses ($p < .006$). However, the MR signal time courses from the right MTL (bottom) showed that HC-correct responses were significantly greater than LC-correct responses ($p < .014$), but there was no difference between HC-incorrect and LC-incorrect responses.

the hypothesis that LC-correct responses are primarily driven by chance responses.

Despite a strong behavioral interaction between confidence and accuracy, statistical activation maps did not show any brain regions with a significant interaction effect ($p < .001$).

DISCUSSION

In this study, we demonstrated that neural activity during encoding of face–name associations is related to the confidence level expressed during a subsequent memory test. We showed a distinct neural representation in the left DLPFC during encoding related to subsequent high confidence, even in instances when confidence and accuracy are dissociated. In contrast, activity in the MTLs showed increased activation only for responses that were both confident and accurate. These results clarify the role of the left prefrontal cortex and MTL in successful memory. Furthermore, the finding that incorrect confident responses showed increased activation in the left prefrontal cortex alone suggests that a subjective feeling of confidence for an incorrect response may be related to incomplete activation of the neural network seen for the formation of successful memory. Our findings may also have implications for situations in which confidence and accuracy are dissociated, such as eyewitness misidentifications and the memory disturbances frequently seen in patients with frontal lobe injuries.

To our knowledge, our study is the first neuroimaging study to investigate encoding as a possible source of confidence–accuracy dissociations. In our paradigm, the majority of high-confidence responses were accurate, and a smaller number of these trials were inaccurate, which represented the particular cases of interest when confidence and accuracy diverge. Instances when subjects have expressed high confidence for an incorrect identification can be viewed as one kind of false memory. A few studies have investigated the origins of false memories during encoding (Gonsalves & Paller, 2000; Mitchell, Dodson, & Schacter, submitted). Gonsalves & Paller (2000) showed that brain potentials during encoding of subsequent false memories were similar to those for true memories. Mitchell et al. (submitted) used event-related fMRI to scan subjects during encoding of trivia statements that were cued as either true or false, and later during a truth judgment task, subjects classified old and new statements. They showed that encod-

ing of cued-false statements that were later classified correctly activated brain regions, including several in the left inferior prefrontal cortex, which had previously been implicated in subsequent recollection. In our study, we showed that a region in the left DLPFC, which was similar to one of the prefrontal regions seen by Mitchell et al., was related to subsequent confidence for both correct and incorrect responses.

Most neuroimaging studies of false recognition have focused on retrieval (e.g., Cabeza, Rao, Wagner, Mayer, & Schacter, 2001; Gonsalves & Paller, 2000; Heun et al., 2000; Schacter, Reiman, et al., 1996; Schacter, Buckner, Koutstaal, Dale, & Rosen, 1997). Many of the fMRI studies during retrieval have also implicated various regions in the prefrontal cortex in false memory. One fMRI study during recognition showed greater activation in the bilateral DLPFC for falsely recognizing a new word compared with correctly rejecting a new word (Cabeza et al., 2001). Other studies have shown that both true and false recognition significantly activated more anterior regions in the left and right prefrontal cortices compared with baseline, but did not find clear differences between the two (Schacter, Reiman, et al., 1996; Schacter et al., 1997). Cabeza et al. (2001) postulated that the prefrontal activation reflects monitoring of the retrieved information, but this notion cannot explain the left DLPFC activation during encoding that we found for HC-incorrect versus LC-incorrect.

Our results suggest that left DLPFC activation during encoding is related to the subjective feeling of high confidence regardless of accuracy because all three contrasts investigating the neural correlates of confidence (all HC vs. all LC, HC-incorrect vs. LC-incorrect, and HC-correct vs. LC-correct) showed significant activation in this region. Very few neuroimaging studies have investigated the subsequent subjective experience of recognition. Maril, Simons, Mitchell, Schwartz, and Schacter (2003) investigated the ''feeling of knowing'' (FOK) that occurs when subjects fail to recall information, but feel that they would recognize the information on a later test. Maril et al. showed that increased activation in a similar left DLPFC and in a left VLPFC region during encoding was associated with subsequent ''know'' and ''FOK'' responses, but also found similar activation for correct ''do not know'' responses. In contrast to our results, which suggest that left DLPFC activation is related to the subjective component of memory, their results suggest that this region is involved in objective accuracy. However, due to inadequate numbers of stimuli, their analyses were confined to correct responses only, so it is possible that incorrect responses could also have shown increased prefrontal activation. In another study that also used confidence ratings in the postscan memory test, Ranganath et al. (2004) demonstrated activation in the DLPFC associated with both familiarity (measured by higher confidence levels) and recollection (measured using a source task). This study also could only investigate the neural correlates for correct responses because the encoding stimuli consisted of single words. However, in light of our results, it also suggests that activation of left inferior prefrontal cortex is related in the subjective feeling of confidence.

Additionally, previous neuroimaging studies during retrieval processes suggest that activation in the left prefrontal cortex may be related to the subjective component of memory. Dobbins et al. (2002) found increased activation in several left prefrontal regions, including the DLPFC, for both correct and incorrect source recognition compared with correct and incorrect item recognition. These activations have been attributed to monitoring operations, which are closely related to subjective experience. Maril et al. (2003), cited earlier, also studied FOK during retrieval and showed a graded response for ''correct know,'' ''correct FOK,'' ''correct do not know,'' and ''incorrect do not know'' responses in the left prefrontal cortex, including the DLPFC, during the retrieval phase. These studies suggest that activation of the left inferior prefrontal cortex during retrieval is related to the subjective memory experience, which also suggests it could be involved in encoding of the subjective memory experience.

The DLPFC was associated with subsequent high confidence for both correct and incorrect responses. The VLPFC, in contrast, showed greater activation in the contrast HC-correct > LC-correct, but not HC-incorrect > LC-incorrect. This provides some evidence that distinct prefrontal subregions may have different roles in encoding. Rypma and D'Esposito (2003) have suggested that activation in the DLPFC is related to subsequent retrieval success (comparing correct to incorrect responses) whereas activation in the VLPFC is related to working memory rehearsal maintenance. However, they only observed DLPFC activation in high memory demand situations, suggesting that the DLPFC is involved in efficiently mediating these working memory functions. Our findings are not consistent with the idea that the DLPFC is primarily involved in subsequent correct identification, since we observed DLPFC activation associated with incorrect confidence responses. However, our observation that VLPFC activation was associated with HC-correct > LC-correct is consistent with the hypothesis that this region is involved in rehearsal.

In addition to activation in the left prefrontal cortex, there were other brain areas that were associated with high confidence for correct responses (HC-correct vs. LC-correct), including fusiform and MTL regions. These brain regions were similar to regions typically reported in subsequent memory studies. This finding is not surprising since most of these studies only report activation in the MTL and left prefrontal regions for events that are remembered with high confidence (Otten, Henson, & Rugg, 2001; Kirchhoff et al., 2000; Wagner

et al., 1998). Our own subsequent memory results showed bilateral hippocampal and left inferior prefrontal activation for face–name pairs that were subsequently remembered with high confidence compared with all incorrect responses (i.e., collapsed across HC-incorrect and LC-incorrect) (Sperling et al., 2003). LC-correct responses are not typically included in subsequent memory designs (Sperling et al., 2003; Otten et al., 2001) and are often assumed to contain correct guesses that occur by chance, which would be functionally equivalent to incorrect responses. Therefore, we expected that the differences in activation for HC-correct versus LC-correct would be similar to HC-correct versus all incorrect. In fact, both contrasts showed greater activity in the right MTL and left prefrontal regions. However, these two contrasts did not show identical patterns of activation. This outcome may reflect the heterogeneity of the LC-correct condition, which may contain both guesses that were correct by chance and responses when subjects were uncertain they had chosen the correct name but did have a slight feeling that it was correct.

Although our SPM interaction analyses did not reveal any regions that demonstrated a significant interaction between confidence and accuracy, statistical comparisons using both SPM contrasts and paired *t* tests on the MR signal response within ROIs showed that there was significantly greater activation in the MTL for HC-correct > LC-correct, but not HC-incorrect > LC-incorrect. This demonstration of differential activation for high confidence for correct compared with incorrect responses suggests that there is indeed an interaction between confidence and accuracy, but we may have had decreased power to detect it.

Our finding that correct confident responses, but not incorrect confident responses, showed MTL activation is consistent with previous studies on MTL function. A number of investigators believe that the primary role of the hippocampus is to bind together distributed representations into a cohesive memory trace (Mayes et al., 2001; Eichenbaum, Otto, & Cohen, 1994; Squire, 1992). Our data are consistent with this idea because successful binding would be indicated by correct identification of the name associated with the face, and our exploratory analyses showed evidence of MTL activation for all correct versus all incorrect. We would therefore expect MTL activation in both the HC-correct and LC-correct condition, but examination of the MR signal response and statistical activation maps show increased MTL activation for HC-correct and not LC-correct responses. As pointed out earlier, the LC-correct condition is presumed to be heterogeneous, with contributions of responses that were correct by chance, which may explain why we did not see the expected increased activation. Additionally, previous subsequent memory studies that defined successful memory based on objective and subjective measures have shown MTL regions are especially important for successful memory (Sperling et al., 2003; Reber et al., 2002; Strange et al., 2002; Otten & Rugg, 2001; Kirchhoff et al., 2000; Wagner et al., 1998; Brewer et al., 1998).

Although each stimulus was presented to the subject for 1.75 sec and there were no significant differences in reaction time for the different conditions, suggesting that time on task did not differ, longer periods of poststimulus fixation at the time of encoding may have aided subsequent memory. The longer periods of jittered fixation, which were necessary in our design for MR signal separation, may have contributed to increased encoding success by allowing extra rehearsal time. In fact, our results did show that there was a very weak correlation between poststimulus fixation period and HC-correct responses ($R = .123$, $p < .01$). Thus, poststimulus fixation period length could have had some contribution to our findings in the MTL. However, Wagner et al. (1998) examined time-on-task and duty-cycle effects by matching the encoding reaction times for high-confidence hits and misses and still demonstrated significant subsequent memory effects in the left prefrontal cortex and MTL. Additionally, Otten et al. (2001) showed more accurate recognition memory for items in a deep encoding task, although the reaction times were shorter than the shallow task, which suggests that time-on-task is not related to depth of processing effects on memory.

Heterogeneity within condition may also have occurred for the other response types in addition to the LC-correct condition. We attempted to control for item familiarity in the postscan test by using distracter names that were paired with other faces during encoding. Thus, other than random trial-to-trial variation, there should not be recognition decisions based on high item strength of the face and target name and low item strength for the distracter name. Additionally, whether the name had previously been presented in the postscan recognition test had a significant effect on confidence level, which may also have contributed to heterogeneity of MR signal response within condition during encoding. We also assumed that the majority of subjects based their decision on a memory for the name associated with the cued face, but it is also possible that in some instances subjects remembered the face associated with the distracter name instead. Thus, it is possible that some of our findings could have arisen from partial encoding and/or retrieval operations.

The cognitive processes involved in the subjective feeling of confidence remains unknown, but behavioral research has offered some possible explanations. Retrospective confidence ratings, such as the one used in our study, seem to be based on factors during both encoding and retrieval (Busey et al., 2000). Behavioral research has shown that the subjective feeling of confidence is related to memory content vividness (Robinson, Johnson, & Robertson, 2000) or to recollection of specific

episodic details (Robinson & Johnson, 1996; Robinson, Johnson, & Herndon, 1997; Robinson et al., 2000). Vividness and encoding of specific details are likely to have origins in encoding. FOK studies have suggested that the FOK may be based on cue familiarity (Koriat & Levy-Sadot, 2001), which could also have origins in encoding. In line with these theories, we do not believe that confidence is a specifically encoded entity that relies on a completely different system, but our results suggest that the neural pattern activated during encoding is related to subsequent confidence. Our results showed that activation in a subset of the regions involved in successful memory, that is increased activation in the left inferior prefrontal cortex in the absence of MTL activation, was related to incorrect confident recognition.

Our findings may be relevant for understanding the neural underpinnings of eyewitness misidentifications. The majority of behavioral studies of this phenomena to date have focused on how factors at the time of retrieval influence the confidence–accuracy (CA) relationship (Wells et al., 2003; Bradfield et al., 2002; Brewer et al., 2002; Perfect et al., 2000; Tomes & Katz, 2000). In one study, active reflection at the time of retrieval on encoding conditions was studied as a factor that impacted the confidence–accuracy relationship (Brewer et al., 2002). Brewer et al. (2002) showed that when participants answered questions about the encoding and retrieval conditions, high confidence was more strongly related to accuracy than when they did not answer questions. Behavioral research has also shown that conditions at the time of encoding influence the confidence–accuracy relationship, and our study provides evidence that it has a neural basis. As mentioned previously, in a videotaped simulated crime, Olsson and Juslin (1999) found that subjects who rated their encoding of the perpetrator as ''holistic'' (deep) rather than ''analytic'' (shallow) showed higher identification accuracy and a stronger correlation between confidence and accuracy. Frost and Weaver (1997), also using a simulated eyewitness event, showed that increasing the time of encoding could eliminate misinformation effects that led to increased misidentifications. Our findings complement these behavioral demonstrations of the importance of encoding factors in the relation between subsequent confidence and accuracy by illuminating the brain regions whose activation is likely relevant to alterations in the confidence–accuracy relationship.

Our results may also be relevant to understanding the kinds of memory failures seen in patients with frontal lobe injuries. Patients with frontal lobe lesions typically show memory deficits on only the most demanding memory tasks (for a review, see Wheeler, Stuss, & Tulving, 1995). Additionally, some patients with frontal lobe damage have shown increased false recognition (Schacter, Curran, Galluccio, Milberg, & Bates, 1996). Because lesions in the frontal lobes do not typically produce a global amnesia, it has been suggested that frontal activations help optimize encoding and retrieval (Fletcher & Henson, 2001). Results from a recent study by Thaiss & Petrides (2003), which examined patients with unilateral frontal lobe lesions, suggest that the frontal cortex is not necessary for retrieval of source memory, but may be involved in generating an effective search strategy for poorly encoded information.

Our results suggest that there are distinct neural patterns of activation during encoding that are related to confidence level during subsequent recognition for face–name pairs. Furthermore, our data indicate that the left prefrontal cortex is involved in subsequent confidence regardless of accuracy. Activation in the right MTL, on the other hand, was related to a combination of subsequent confidence and accuracy. Most interestingly, our data suggest that there may be a neural explanation for the dissociation between confidence and accuracy, seen in eyewitness misidentifications and in patients with frontal damage who exhibit memory impairment. Our data may help explain the origins of confidence–accuracy dissociations, suggesting that the feeling of high confidence in the absence of accuracy may be related to a partial activation of the network required for successful memory formation.

METHODS

Subjects

Sixteen right-handed, healthy young subjects (ages 20–33, 9 women and 7 men) participated in this study. All subjects were free from psychiatric and neurologic illness, and none were taking medications with known central nervous system effects. All subjects were screened for contraindications to MRI. Each subject provided written informed consent in a manner approved by the Human Research Committee at Massachusetts General Hospital, Boston, MA.

Cognitive Task

Stimuli consisted of face–name pairs composed of digital color photographs of unfamiliar faces paired with fictional first names. Each face was presented against a black background with the first name printed in white underneath it. There were approximately equal numbers of men and women and equal numbers of young, middle-aged, and older faces. Approximately one-third of the faces were non-White faces. First names were assigned based on lists of popular first names by decade obtained from the Internet. Across five encoding runs, 455 face–name pairs were presented for 1.75 sec each and were randomly intermixed with brief periods of focused fixation on a white crosshair. Each face–name stimulus was followed by 0.25 sec of fixation, which was then followed by periods of visual fixation of varying lengths (0–10 sec, mean fixation length = 2.84 sec).

Each run contained, on average, 36.8 fixation periods of varying length. Fixation trials were intermixed with stimuli using a jittered design and optimized using OptSeq (http://surfer.nmr.mgh.harvard.edu). Stimuli were presented using MacStim 2.5 on a Macintosh G3 computer and viewed using a rear projection system. There was approximately 1 min between runs. The entire encoding session lasted approximately 30 min.

We used an intentional encoding task, in which each subject was instructed to try to remember the name associated with each face for later testing. Subjects were also instructed to indicate by pressing a button using their right hand whether they subjectively thought the name "fit" the face. The types of button response ("fit" vs. "not-fit") and the response times were recorded using MacStim. This task was chosen to aid in the associative memory process and to ensure that subjects attended to both the name and the face.

Approximately 15 min after the scanning session, subjects completed a postscan recognition test of all 455 face–name pairs shown during scanning. The test was a forced-choice recognition task in which subjects viewed each face seen during encoding paired with the correct name and an incorrect name. To hold familiarity levels for the names constant, the incorrect name was a name that was previously paired with a different face during encoding. Subjects indicated which name was the correct name (left or right) by pressing one of two buttons on the laptop computer and then indicated their confidence level that they had chosen the correct name (high or low) by pressing one of two other buttons. The recognition test was self-paced and subjects had a maximum of 20 sec to make each decision. Subjects were not told that each name had presented with only one face, that each face was presented with a unique first name, that all faces and names had been seen during encoding, nor were they told that the correct name would indeed appear on the screen with the face. Subjects were told to indicate "high confidence" if they were sure they had chosen the correct name and "low confidence" if they guessed or were not sure of the correct name. On the memory test, the correct name appeared with equal frequency on the left and the right.

Analyses of behavioral data were conducted using SPSS (Chicago, IL). Chi-square analyses were performed on each individual subject to determine whether they performed about chance levels for accurate (all correct) and for confident and accurate (HC-correct) responses. Objective accuracy was based only on whether the subjects correctly identified the name associated with the face. Repeated measures ANOVAs were used to determine any differences in reaction times and number of trials for each response type. Paired t tests were used to determine differences related to the encoding task about whether the name "fit" the face. Results were considered significant at $p < .05$, two-tailed.

Functional Imaging

The fMRI data were collected on a Siemens 3T Allegra scanner (Siemens Medical Systems, Iselin, NJ) with a three-axis gradient head coil using a gradient-echo echo-planar pulse sequence (TR = 2000 msec, TE = 30 msec, 28 oblique coronal slices perpendicular to the anterior commissure–posterior commissure line, 5-mm slices, skip 1 mm). Each functional run consisted of 140 time points and lasted 4 min and 40 sec.

Data Analysis

The fMRI data were preprocessed and analyzed using SPM99 (Wellcome Department on Cognitive Neurology) for Matlab (Mathworks, Natick, MA). No slice-timing correction wasapplied. Images were realigned to correct for motion using sinc interpolation. The data were then spatially normalized to an EPI template based on the MNI1305 stereotactic space and then spatially smoothed using an 8-mm full width half maximum isotropic Gaussian kernel.

Encoding trials were categorized based on the four possible response types from the postscan recognition task: HC-correct, LC-correct, HC-incorrect, and LC-incorrect. Subjects were required to have at least 10 events for each trial type to be included in the analysis. Data were analyzed according to a mixed-effects general linear model in SPM. First, data were analyzed at the subject level, with each run treated as a time series, and modeled with the canonical hemodynamic response function only. To eliminate low frequency noise, data were analyzed using a high-pass filter of 70 sec. At the second step, data were averaged together, treating each subject as a random effect. We assessed the main effect of confidence by collapsing the responses across accuracy with the contrast all HC versus all LC. We also generated contrasts of confidence for incorrect responses (HC-incorrect vs. LC-incorrect) and confidence for correct responses (HC-correct vs. LC-correct). We assessed the main effect of accuracy for choosing the name associated with the face with the contrast all correct versus all incorrect. We also generated contrasts of accuracy within high-confidence (HC-correct vs. HC-incorrect) and within low-confidence responses (LC-correct vs. LC-incorrect. We also generated contrasts for the interaction between confidence and accuracy. For contrasts, "versus" is used when referring to differences in activation in either direction and "<" and ">" are used to denote the specific direction for the contrast. Statistical maps were thresholded at a significance level of $p < .001$ (one-tailed, uncorrected) with an extent threshold of 20 contiguous voxels. In regions for which we had a priori hypotheses about their involvement or to minimize the probability of false negatives, we used a less stringent threshold of $p < .01$ in exploratory analyses. MNI coordinates are reported for the

regions of activation. The MTL was treated as a whole unit that encompasses both the hippocampus and entorhinal cortex because our data were spatially smoothed, thus making it difficult to definitively distinguish subregions.

To determine which brain regions were involved in high confidence regardless of accuracy, we used a masking approach (Cabeza et al., 2002) to determine overlapping brain regions that were active in both HC-incorrect versus LC-incorrect and HC-correct versus LC-correct contrasts. Masks for the appropriate contrasts were multiplied together, yielding a new mask that contained voxels that were active for each contrast. Because the joint probability that the same region is significantly active in both contrasts was less that the probability the region would be active in one contrast, we used a threshold of $p < .01$ for the individual contrasts used to make the masks.

Percent signal modulation for functionally defined ROIs was determined using the SPM ROI toolbox (http://spm-toolbox.sourceforge.net). ROIs were defined by including all significant voxels within a 6-mm radius of a peak voxel from a contrast of interest. Graphs were generated using oversampled data, with data points plotted every 0.5 sec. Differences in the MR Signal response were tested between conditions at the peak of 4 or 6 sec using SPSS.

Acknowledgments

We thank Mary Foley, Jennifer Holmes, and Larry White for help with scan acquisition. This work was supported by NINDS (K23-NS02189; R. A. S.), NIMH (MH60941; D. L. S.), and NIA (P01-AG-04953; M. S. A.) and Harvard Center for Neurodegeneration and Repair.

Reprint requests should be sent to Elizabeth F. Chua, Center for Neurocognitive Studies, Department of Neurology, Brigham and Women's Hospital, 221 Longwood Ave., Boston, MA 02115, or via e-mail: echua@rics.bwh.harvard.edu.

The data reported in this experiment have been deposited in the fMRI Data Center (http://www.fmridc.org). The accession number is 2-2004-115W1.

REFERENCES

Belli, R. F., Lindsay, D. S., Gales, M. S., & McCarthy, T. T. (1994). Memory impairment and source misattribution in postevent misinformation experiments with short retention intervals. *Memory and Cognition, 22,* 40–54.

Bornstein, B. H., & Zickafoose, D. J. (1999). ''I know I know it, I know I saw it'': The stability of the confidence–accuracy relationship across domains. *Journal of Experimental Psychology: Applied, 5,* 76–88.

Bradfield, A. L., Wells, G. L., & Olson, E. A. (2002). The damaging effect of confirming feedback on the relation between eyewitness certainty and identification accuracy. *Journal of Applied Psychology, 87,* 112–120.

Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1998). Making memories: Brain activity that predicts how well visual experience will be remembered [see comments]. *Science, 281,* 1185–1187.

Brewer, N., Keast, A., & Rishworth, A. (2002). The confidence–accuracy relationship in eyewitness identification: The effects of reflection and disconfirmation on correlation and calibration. *Journal of Experimental Psychology: Applied, 8,* 44–56.

Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence–accuracy relation in recognition memory. *Psychonomic Bulletin & Review, 7,* 26–48.

Cabeza, R., Dolcos, F., Graham, R., & Nyberg, L. (2002). Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *Neuroimage, 16,* 317–330.

Cabeza, R., Rao, S. M., Wagner, A. D., Mayer, A. R., & Schacter, D. L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Sciences, U.S.A., 98,* 4805–4810.

Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Mapping, 5,* 329–340.

Daselaar, S. M., Veltman, D. J., Rombouts, S. A. R. B., Raaijmakers, J. G. W., & Jonker, C. (2003). Neuroanatomical correlates of episodic encoding and retrieval in young and elderly subjects. *Brain, 126,* 43–56.

Dobbins, I., Foley, H., Schacter, D., & Wagner, A. (2002). Executive control during episodic retrieval. Multiple prefrontal processes subserve source memory. *Neuron, 35,* 989.

Eichenbaum, H. B., Otto, T., & Cohen, N. J. (1994). Two functional components of the hippocampal memory system. *Behavioral and Brain Sciences, 17,* 449–518.

Fletcher, P. C., & Henson, R. N. (2001). Frontal lobes and human memory: Insights from functional neuroimaging. *Brain, 124,* 849–881.

Frost, P., & Weaver, C. A., III. (1997). Overcoming misinformation effects in eyewitness memory: Effects of encoding time and event cues. *Memory, 5,* 725–740.

Gonsalves, B., & Paller, K. A. (2000). Neural events that underlie remembering something that never happened. *Nature Neuroscience, 3,* 1316–1321.

Henson, R. N., Rugg, M. D., Shallice, T., & Dolan, R. J. (2000). Confidence in recognition memory for words: Dissociating right prefrontal roles in episodic retrieval. *Journal of Cognitive Neuroscience, 12,* 913–923.

Henson, R. N., Rugg, M. D., Shallice, T., Josephs, O., & Dolan, R. J. (1999). Recollection and familiarity in recognition memory: An event-related functional magnetic resonance imaging study. *Journal of Neuroscience, 19,* 3962–3972.

Heun, R., Jessen, F., Klose, U., Erb, M., Granath, D. O., & Grodd, W. (2000). Response-related fMRI analysis during encoding and retrieval revealed differences in cerebral activation by retrieval success. *Psychiatry Research, 99,* 137–150.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In F. I. Craik (Ed.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–423). Mahwah, NJ: Erlbaum.

Kirchhoff, B. A., Wagner, A. D., Maril, A., & Stern, C. E. (2000). Prefrontal–temporal circuitry for episodic encoding and subsequent memory. *Journal of Neuroscience, 20,* 6173–6180.

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 34–53.

Lindsay, D. S., Read, J. D., & Sharma, K. (1998). Accuracy and confidence in person identification: The relationship is strong when witnessing conditions vary widely. *Psychological Science, 9,* 215–218.

Maril, A., Simons, J. S., Mitchell, J. P., Schwartz, B. L., & Schacter, D. L. (2003). Feeling-of-knowing in episodic memory: An event-related fMRI study. *Neuroimage, 18,* 827–836.

Mayes, A. R., Isaac, C. L., Holdstock, J. S., Hunkin, N. M., Montaldi, D., & Downes, J. J. (2001). Memory for single items, word pairs, and temporal order of different kinds in a patient with selective hippocampal lesions. *Cognitive Neuropsychology, 18,* 97–123.

Mitchell, J. P., Dodson, C. S., & Schacter, D. L. (submitted). Counteracting misattribution: An event-related fMRI study of illusory truth.

Olsson, N., & Juslin, P. (1999). Can self-reported encoding strategy and recognition skill be diagnostic of performance in eyewitness identifications? *Journal of Applied Psychology, 84,* 42–49.

Otten, L. J., Henson, R. N., & Rugg, M. D. (2001). Depth of processing effects on neural correlates of memory encoding: Relationship between findings from across- and within-task comparisons. *Brain, 124,* 399–412.

Otten, L. J., & Rugg, M. D. (2001). Task-dependency of the neural correlates of episodic encoding as measured by fMRI. *Cerebral Cortex, 11,* 1150–1160.

Perfect, T. J., Hollins, T. S., & Hunt, A. L. (2000). Practice and feedback effects on the confidence–accuracy relation in eyewitness memory. *Memory, 8,* 235–244.

Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia, 42,* 2–13.

Reber, P. J., Siwiec, R. M., Gitelman, D. R., Parrish, T. B., Mesulam, M. M., Paller, K. A. (2002). Neural correlates of successful encoding identified using functional magnetic resonance imaging. *Journal of Neuroscience, 22,* 9541–9548.

Robinson, M. D., & Johnson, J. T. (1996). Recall memory, recognition memory, and the eyewitness confidence–accuracy correlation. *Journal of Applied Psychology, 81,* 587–594.

Robinson, M. D., Johnson, J. T., & Herndon, F. (1997). Reaction time and assessments of cognitive effort as predictors of eyewitness memory accuracy and confidence. *Journal of Applied Psychology, 82,* 416–425.

Robinson, M. D., Johnson, J. T., & Robertson, D. A. (2000). Process versus content in eyewitness metamemory monitoring. *Journal of Experimental Psychology: Applied, 6,* 207–221.

Rypma, B., & D'Esposito, M. (2003). A subsequent-memory effect in dorsolateral prefrontal cortex. *Brain Research Cognitive Brain Research, 16,* 162–166.

Schacter, D. L., Buckner, R. L., Koutstaal, W., Dale, A. M., & Rosen, B. R. (1997). Late onset of anterior prefrontal activity during true and false recognition: An event-related fMRI study. *Neuroimage, 6,* 259–269.

Schacter, D. L., Curran, T., Galluccio, L., Milberg, W. P., & Bates, J. F. (1996). False recognition and the right frontal lobe: A case study. *Neuropsychologia, 34,* 793–808.

Schacter, D. L., Reiman, E., Curran, T., Yun, L. S., Bandy, D., McDermott, K. B. (1996). Neuroanatomical correlates of veridical and illusory recognition memory: Evidence from positron emission tomography. *Neuron, 17,* 267–274.

Shaw, J. S., III, & Zerr, T. K. (2003). Extra effort during memory retrieval may be associated with increases in eyewitness confidence. *Law and Human Behavior, 27,* 315–329.

Sperling, R., Bates, J., Cocchiarella, A., Schacter, D., Rosen, B., & Albert, M. (2001). Encoding novel face–name associations: A functional MRI study. *Human Brain Mapping, 14,* 129–139.

Sperling, R., Chua, E., Cocchiarella, A., Rand-Giovannetti, E., Poldrack, R., Schacter, D. L. (2003). Putting names to faces: Successful encoding of associative memories activates the anterior hippocampal formation. *Neuroimage, 20,* 1400–1410.

Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence–accuracy relation in eyewitness identification studies. *Psychological Bulletin, 118,* 315–327.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans [published erratum appears in Psychol Rev 1992 Jul;99(3):582]. *Psychological Review, 99,* 195–231.

Strange, B. A., Otten, L. J., Josephs, O., Rugg, M. D., & Dolan, R. J. (2002). Dissociable human perirhinal, hippocampal, and parahippocampal roles during verbal encoding. *Journal of Neuroscience, 22,* 523–528.

Stretch, V., & Wixted, J. T. (1998). Decision rules for recognition memory confidence judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1397–1410.

Thaiss, L., & Petrides, M. (2003). Source versus content memory in patients with a unilateral frontal cortex or a temporal lobe excision. *Brain, 126,* 1112–1126.

Tomes, J. L., & Katz, A. N. (2000). Confidence–accuracy relations for real and suggested events. *Memory, 8,* 273–283.

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity [see comments]. *Science, 281,* 1188–1191.

Weber, N., & Brewer, N. (2003). The effect of judgment type and confidence scale on confidence–accuracy calibration in face recognition. *Journal of Applied Psychology, 88,* 490–499.

Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology, 54,* 277–295.

Wells, G. L., Olson, E. A., & Charman, S. D. (2002). The confidence of eyewitnesses in their identifications from lineups. *Current Directions in Psychological Science, 11,* 151–154.

Wells, G. L., Olson, E. A., & Charman, S. D. (2003). Distorted retrospective eyewitness reports as functions of feedback and delay. *Journal of Experimental Psychology: Applied, 9,* 42–52.

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1995). Frontal lobe damage produces episodic memory impairment. *Journal of International Neuropsychological Society, 1,* 525–536.

This article has been cited by:

1. Kelly Sullivan Giovanello, David Schnyer, Mieke Verfaellie. 2009. Distinct hippocampal regions make unique contributions to relational memory. *Hippocampus* **19**:2, 111-117. [CrossRef]
2. Sara Kvidera, Wilma Koutstaal. 2008. Confidence and decision type under matched stimulus conditions: overconfidence in perceptual but not conceptual decisions. *Journal of Behavioral Decision Making* **21**:3, 253-281. [CrossRef]
3. P. Willner. 2008. Clarification of the memory artefact in the assessment of suggestibility. *Journal of Intellectual Disability Research* **52**:4, 318-326. [CrossRef]
4. Elizabeth F. Chua, Daniel L. Schacter, Erin Rand-Giovannetti, Reisa A. Sperling. 2007. Evidence for a specific role of the anterior hippocampal region in successful associative encoding. *Hippocampus* **17**:11, 1071-1080. [CrossRef]
5. G.F. Wagstaff, J. Cole, J. Wheatcroft, M. Marshall, I. Barsby. 2007. A componential approach to hypnotic memory facilitation: focused meditation, context reinstatement and eye movements. *Contemporary Hypnosis* **24**:3, 97-108. [CrossRef]
6. Yun-Ching Kao, Emily S Davis, John D E Gabrieli. 2006. Neural correlates of actual and predicted memory formation. *Nature Neuroscience* **8**:12, 1776-1783. [CrossRef]
7. 2005. Current Awareness in NMR in Biomedicine. *NMR in Biomedicine* **18**:2, 135-142. [CrossRef]